



PCT

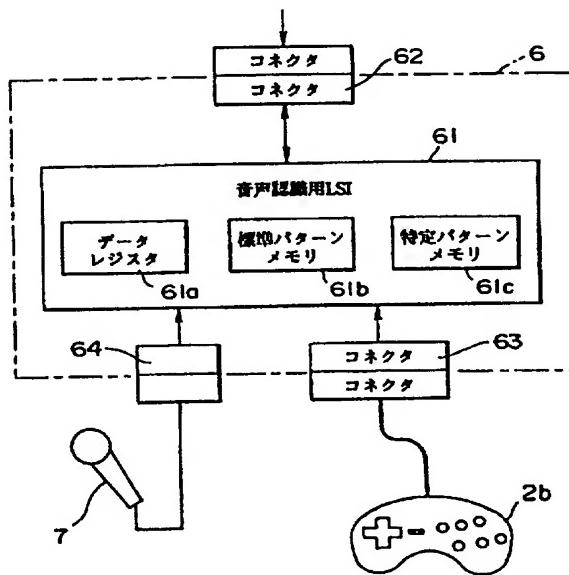
特許協力条約に基づいて公開された国際出願

(51) 国際特許分類6 A63F 9/22	A1	(11) 国際公開番号 WO98/02223
		(43) 国際公開日 1998年1月22日(22.01.98)
(21) 国際出願番号 PCT/JP97/02421		
(22) 国際出願日 1997年7月11日(11.07.97)		
(30) 優先権データ 特願平8/182495 1996年7月11日(11.07.96)	JP	(81) 指定国 CN, JP, KR, US, 欧州特許 (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).
(71) 出願人 (米国を除くすべての指定国について) 株式会社 セガ・エンタープライゼス (SEGA ENTERPRISES, LTD.)[JP/JP] 〒144 東京都大田区羽田1丁目2番12号 Tokyo, (JP)		添付公開書類 国際調査報告書 補正書
(72) 発明者 ; および (75) 発明者 / 出願人 (米国についてのみ) 中島拓也(NAKAJIMA, Takuya)[JP/JP] 浜田和彦(HAMADA, Kazuhiko)[JP/JP] 〒144 東京都大田区羽田1丁目2番12号 株式会社 セガ・エンタープライゼス内 Tokyo, (JP)		
(74) 代理人 弁理士 稲葉良幸, 外(JNABA, Yoshiyuki et al.) 〒105 東京都港区虎ノ門3丁目5番1号 37森ビル803号室 TMI総合法律事務所 Tokyo, (JP)		

(54) Title: VOICE RECOGNIZER, VOICE RECOGNIZING METHOD AND GAME MACHINE USING THEM

(54) 発明の名称 音声認識装置、音声認識方法及びこれを用いたゲーム機

61 ... voice recognition LSI
61a ... data register
61b ... reference pattern memory
61c ... specific pattern memory
62, 63 ... connector



(57) Abstract

A voice recognizer and a voice recognizing method which are used as peripheral apparatuses are provided in order to enjoy a game through speech. A voice input means (7), a game machine control unit (2b) and a voice recognition unit (61) which compares the voice signals outputted from the voice input means (7) with data in predetermined voice recognition dictionaries (61b and 61c) to recognize the voice of a player and generates control signals concerning the game in accordance with the recognition result and the output signals of the control unit (2b).

(57) 要約

音声によりゲームを楽しむために、周辺機器として用いられる音声認識装置及び音声認識方法を提供する。音声入力手段7と、ゲーム機用の制御部2bと、音声入力手段7の音声信号出力と予め定められた音声認識用辞書61b、61cのデータとを比較することにより遊戯者の音声を認識するとともに、認識結果と制御部2bの出力信号とに基づきゲームに関する制御信号を生成する音声認識部61とを備える。

参考情報

PCTに基づいて公開される国際出願のパンフレット第一頁に記載されたPCT加盟国を同定するために使用されるコード

AL アルバニア	ES スペイン	LR リベリア	SG シンガポール
AM アルメニア	FI フィンランド	LS レソト	SI スロヴェニア
AT オーストリア	FR フランス	LT リトアニア	SK スロヴァキア共和国
AU オーストラリア	GA ガボン	LU ルクセンブルグ	SL シエラレオネ
AZ アゼルバイジャン	GB 英国	LV ラトヴィア	SN セネガル
BA ボスニア・ヘルツェゴビナ	GE グルジア	MC モナコ	SZ スウェーデン
BB バルバドス	GH ガーナ	MD モルドバ共和国	TD チャード
BE ベルギー	GM ガンビア	MG マダガスカル	TG トーゴ
BF ブルキナ・ファソ	GN ギニア	MK マケドニア旧ユーゴス	TJ タジキスタン
BG ブルガリア	GR ギリシャ	LA ラヴィア共和国	TM トルクメニスタン
BJ ベナン	HU ハンガリー	ML マリ	TR トルコ
BR ブラジル	ID インドネシア	MN モンゴル	TT トリニダード・トバゴ
BY ベラルーシ	IE アイルランド	MR モーリタニア	UA ウクライナ
CA カナダ	IL イスラエル	MW マラウイ	UG ウガンダ
CF 中央アフリカ共和国	IS アイスランド	MX メキシコ	US 米国
CG コンゴー	IT イタリア	NE ニジェール	UZ ウズベキスタン
CH スイス	JP 日本	NL オランダ	VN ヴィエトナム
CIM コート・ジボアール	KE ケニア	NO ノルウェー	YU ユーゴスラビア
CM カメルーン	KG キルギスタン	NZ ニュージーランド	ZW ジンバブエ
CN 中国	KP 朝鮮民主主義人民共和国	PL ポーランド	
CU キューバ	KR 大韓民国	PT ポルトガル	
CZ チェコ共和国	KZ カザフスタン	RO ルーマニア	
DE ドイツ	LC セントルシア	RU ロシア連邦	
DK デンマーク	LI リヒテンシュタイン	SD スーダン	
EE エストニア	LK スリランカ	SE スウェーデン	

明細書

音声認識装置、音声認識方法及びこれを用いたゲーム機

技術分野

この発明は、入力された音声を認識する音声認識装置、音声認識方法及びこれを用いたゲーム機に関するものである。

背景技術

近年のコンピュータ技術の進歩に伴い、コンピュータグラフィックス技術を用いたビデオゲーム機が広く利用されるようになってきた。近年多種多様なゲームソフトが開発されているため、この種の家庭用のビデオゲーム装置は娯楽として非常に人気が高い。この種のビデオゲーム機ではコントローラのボタンを操作することにより、キャラクタを動かしたり、ゲーム画面を切り替えたりしている。

このように従来のビデオゲーム機において、ゲームの指示はすべて遊戯者がコントローラのスイッチを操作して行うようになっていた。

ところで、人間にとて最も自然なコミュニケーション手段は音声である。このため、人間の音声を認識し、この認識結果に基づき機械を制御しようとする試みが各種行われている。例えば、車やロボット等を人間の音声により制御しようとしたり、音声による指示により各種電子装置を制御しようとしていた。

しかし、従来の音声認識装置をビデオゲーム機に適用しようとした場合、一般的な電子装置の場合と異なりゲーム機の特性にあった装置でなければならず、その

まま適用はできない。例えば、ゲーム機のコントローラのボタンを押すときに、単に1回押すだけでなく、長く押したりくり返し押すことによりキャラクタの動作の程度を変えたり、他のボタンと同時に押すことにより特別な動作をさせたりしなければならない。また、レスポンスの良いことも重要である。一方、ゲームの種類によっては認識の精度はあまり重要とならず、却って、なかなか認識できないこと自体がゲームの面白さに通じることもある。

要するに、遊戯者がゲームを楽しむためには、従来の音声認識装置の場合とは異なる特有の性能・機能を備える必要がある。この種の要求を満たす音声認識装置は、従来、知られていなかった。

この発明は、係る課題を解決するためになされたものであり、ゲームに適する音声認識装置及び音声認識方法を提供することを目的とする。

また、この発明は、前記音声認識装置あるいは音声認識方法を用いたゲーム機を提供することを目的とする。

発明の開示

この発明に係る音声認識装置は、ゲーム機の周辺装置として用いられる音声認識装置であって、音声入力手段と、前記音声入力手段の音声信号出力と予め定められた音声認識用辞書のデータとを比較することにより遊戯者の音声を認識するとともに、認識結果に基づきゲームに関する制御信号を生成する音声認識部とを備えるものである。

この発明に係る音声認識装置は、前記音声認識部は、前記音声認識用辞書として、不特定話者用に予め設定された不特定話者音声認識用辞書及び遊戯者により設定される特定話者音声認識用辞書を備え、動作の初期状態において前記不特定

話者音声認識用辞書を選択するものである。

この発明に係る音声認識装置は、前記音声認識部は、複数の遊戯者それぞれに対応する複数の特定話者音声認識用辞書を備え、遊戯者の操作によりこれら特定話者音声認識用辞書から認識処理に使用する辞書を選択するものである。

この発明に係る音声認識装置は、前記音声認識部に接続されるゲーム機用の制御部を備え、前記音声認識部は、音声の認識結果と前記制御部の制御信号とを組み合わせてゲームに関する制御信号を生成するものである。

この発明に係る音声認識装置は、前記制御部は、通常の動作をさせる制御信号を出力し、前記音声認識部は、特殊な動作をさせる制御信号を生成するものである。

通常の動作とは、遊戯者が通常行うゲームにおける動作（例えば、キックする、ジャンプする等）である。特殊な動作とは、前記制御部の信号と前記音声信号との組み合わせにより可能になる特有の動作である。例えば、格闘技ゲームにおいて、大きな声+Aボタン操作により、そのときの同じ技が繰り返すような動作や、複数のボタンを同時に押した場合にのみ出せる特定の必殺技である。

この発明に係る音声認識装置は、前記音声認識部は、前記音声入力手段の音声信号出力と前記音声認識用辞書の内容との類似の程度を示す類似度及び前記音声信号の状態を示す値を出力するとともに、前記音声認識部は、ゲームにおける動作の種類を選択するために前記類似度に基づき音声認識を行う第1の動作モードと、動作の状態を設定するために音声認識を行わずに前記音声信号の状態を測定する第2の動作モードとを備えるものである。

この発明に係る音声認識装置は、前記第2のモードにおいて、前記音声信号の状態を、信号の一部あるいは全体の平均音量レベルとし、この平均音量レベルに

に基づき動作の状態を設定するものである。

この発明に係る音声認識装置は、前記第2のモードにおいて、前記音声信号の状態を、信号のピーク音量レベルとし、このピーク音量レベルに基づき動作の状態を設定するものである。

この発明に係る音声認識装置は、前記第2のモードにおいて、前記音声信号の状態を、信号の立ち上がり時間とし、この立ち上がり時間に基づき動作の状態を設定するものである。

この発明に係る音声認識装置は、前記第2のモードにおいて、前記音声信号の状態を、信号の持続時間とし、この持続時間に基づき動作の状態を設定するものである。

この発明に係る音声認識装置は、前記第2のモードにおいて、前記音声信号の状態を、音声の種類とし、この音声の種類に基づき動作の状態を設定するものである。

この発明に係る音声認識装置は、前記音声認識部は、前記音声入力手段の音声信号出力と前記音声認識用辞書の内容との類似の程度を示す類似度及びその音量レベルを出力し、この音量レベルを予め定められたリジェクトレベルに基づき評価し、評価結果に基づき前記音声認識部の認識結果をリジェクトするものである。

この発明に係る音声認識装置は、前記リジェクトレベルをゲームの種類あるいはゲームステージごとに設定するものである。

この発明に係る音声認識装置は、前記音声認識部は、前記音声入力手段の音声信号出力と前記音声認識用辞書の内容との類似の程度を示す類似度に基づき認識するとともに、認識結果に基づき生成された制御信号によるゲームにおける動作の状態を前記類似度に応じて変化させるものである。

この発明に係るゲーム機は、制御用周辺機器として、前記音声認識装置を備えるものである。

この発明に係る音声認識方法は、音声信号を受ける第1のステップと、前記音声信号と予め定められた音声認識用辞書のデータとを比較することにより遊戯者の音声を認識する第2のステップと、前記第2のステップの認識結果に基づきゲームに関する制御信号を生成する第3のステップとを備えるものである。

この発明に係る音声認識方法は、ゲーム機用の制御部から制御信号を受ける第4のステップを備え、前記第3のステップにおいて、前記第2のステップによる音声の認識結果と前記第4のステップによる制御信号とを組み合わせてゲームに関する制御信号を生成するものである。

図面の簡単な説明

図1は、この発明の実施の形態1に係るビデオゲーム機を用いたゲームシステムの外観図であり、図2は、この発明の実施の形態1に係るビデオゲーム機の機能ブロック図であり、図3は、この発明の実施の形態1に係る音声認識部の機能ブロック図であり、図4は、この発明の実施の形態1に係るビデオゲーム機を用いたゲームシステムの使用状態の説明図であり、図5は、この発明の実施の形態1に係る音声認識部の動作を説明するためのフローチャートであり、図6は、この発明の実施の形態1に係る音声認識部の動作を説明するための音声波形の概略図であり、図7は、この発明の実施の形態1に係る音声認識部の動作を説明するための音声波形の概略図である。

発明を実施するための最良の形態

形態 1.

図 1 はこの発明の実施の形態に係る音声認識装置を用いたビデオゲーム機の外観図である。この図において、ビデオゲーム機本体 1 は略箱型をなし、その内部にはゲーム処理用の基板等が設けられている。また、ビデオゲーム機本体 1 の前面には、2つのコネクタ 2 a が設けられており、これらのうちの一方のコネクタ 2 a にはゲーム操作用の P A D 2 b がケーブル 2 c を介して接続されている。他方のコネクタ 2 a には音声認識部 6 が接続されている。音声認識部 6 には、P A D 2 b とともに遊戯者の音声を聞き取るためのマイク 7 が接続されている。2人の遊戯者がゲームを楽しむ場合には、2つの P A D 2 b が使用される。図 1 のマイク 7 はヘッドセット型であるが、他の形式のマイクでもよい。なお、2つのコネクタ 2 a の両方に音声認識部 6 を接続することもできる。

ビデオゲーム機本体 1 の上部には、ゲームプログラム、音声認識動作が記憶された記憶媒体としての R O M カートリッジ接続用のカートリッジ I / F 1 a、C D - R O M 読み取り用の C D - R O M ドライブ 1 b が設けられている。ビデオゲーム機本体 1 の背面には、図示されていないが、ビデオ出力端子およびオーディオ出力端子が設けられている。このビデオ出力端子はケーブル 4 a を介して T V 受像機 5 のビデオ入力端子に接続されるとともに、オーディオ出力端子はケーブル 4 b を介して T V 受像機 5 のオーディオ入力端子に接続されている。このようなビデオゲーム機において、ユーザが P A D 2 b を操作することにより、T V 受像機 5 に映し出された画面を見ながらゲームを行うことができる。

図 2 は本実施の形態に係る T V ゲーム機の概要を表すブロック図である。なお、

図2において、音声認識部6及びマイク7の表示は省略している。この画像処理装置は、装置全体の制御を行うCPUブロック10、ゲーム画面の表示制御を行うビデオブロック11、効果音等を生成するサウンドブロック12、CD-ROMの読み出しを行うサブシステム13等により構成されている。

CPUブロック10は、SCU(System Control Unit)100、メインCPU101、RAM102、ROM103、カートリッジI/F1a、サブCPU104、CPUバス103等により構成されている。メインCPU101は、装置全体の制御を行うものである。このメインCPU101は、内部にDSP(Digital Signal Processor)と同様の演算機能を備え、アプリケーションソフトを高速に実行可能である。RAM102は、メインCPU101のワークエリアとして使用されるものである。ROM103には、初期化処理用のイニシャルプログラム等が書き込まれている。SCU100は、バス105、106、107を制御することにより、メインCPU101、VDP120、130、DSP140、CPU141等の間におけるデータ入出力を円滑に行うものである。また、SCU100は、内部にDMAコントローラを備え、ゲーム中のスプライトデータをビデオブロック11内のVRAMに転送することができる。これにより、ゲーム等のアプリケーションソフトを高速に実行することが可能である。カートリッジI/F1aは、ROMカートリッジの形態で供給されるアプリケーションソフトを入力するためのものである。

サブCPU104は、SMPC(System Manager & Peripheral Control)と呼ばれるもので、メインCPU101からの要求に応じて、PAD2bからペリフェラルデータをコネクタ2aを介して収集する機能等を備えている。メインCPU101はサブCPU104から受け取ったペリフェラルデータに基づき、例え

ばゲーム画面中の野手を移動させる等の処理を行うものである。コネクタ2aには、P A D、ジョイスティック、キーボード等のうちの任意のペリフェラルが接続可能である。サブC P U 1 0 4は、コネクタ2a（本体側端子）に接続されたペリフェラルの種類を自動的に認識し、ペリフェラルの種類に応じた通信方式に従いペリフェラルデータ等を収集する機能を備えている。

ビデオブロック11は、ビデオゲームのポリゴンデータから成るキャラクタ等の描画を行うV D P（Video Display Processor）120、背景画面の描画、ポリゴン画像データおよび背景画像の合成、クリッピング処理等を行うV D P 1 3 0とを備えている。V D P 1 2 0はV R A M 1 2 1およびフレームバッファ122、123に接続されている。ビデオゲーム機のキャラクタを表すポリゴンの描画データはメインC P U 1 0 1からS C U 1 0 0を介してV D P 1 2 0に送られ、V R A M 1 2 1に書き込まれる。V R A M 1 2 1に書き込まれた描画データは、例えば、16または8ビット/pixelの形式で描画用のフレームバッファ122または123に描画される。描画されたフレームバッファ122または123のデータはV D P 1 3 0に送られる。描画を制御する情報は、メインC P U 1 0 1からS C U 1 0 0を介してV D P 1 2 0に与えられる。そして、V D P 1 2 0は、この指示に従い描画処理を実行する。

V D P 1 3 0はV R A M 1 3 1に接続され、V D P 1 3 0から出力された画像データはメモリ132を介してエンコーダ160に出力される構成となっている。エンコーダ160は、この画像データに同期信号等を付加することにより映像信号を生成し、T V受像機5に出力する。これにより、T V受像機5に所定のゲームの画面が表示される。

サウンドブロック12は、P C M方式あるいはF M方式に従い音声合成を行う

DSP140と、このDSP140の制御等を行うCPU141により構成されている。DSP140により生成された音声データは、D/Aコンバータ170により2チャネルの信号に変換された後にスピーカ5bに出力される。

サブシステム13は、CD-ROMドライブ1b、CD I/F180、CPU181、MPEG AUDIO182、MPEG VIDEO183等により構成されている。このサブシステム13は、CD-ROMの形態で供給されるアプリケーションソフトの読み込み、動画の再生等を行う機能を備えている。CD-ROMドライブ1bはCD-ROMからデータを読み取るものである。CPU181は、CD-ROMドライブ1bの制御、読み取られたデータの誤り訂正等の処理を行うものである。CD-ROMから読み取られたデータは、CD I/F180、バス106、SCU100を介してメインCPU101に供給され、アプリケーションソフトとして利用される。また、MPEG AUDIO182、MPEG VIDEO183は、MPEG規格 (Motion Picture Expert Group) により圧縮されたデータを復元するデバイスである。これらのMPEG AUDIO182、MPEG VIDEO183を用いてCD-ROMに書き込まれたMPEG圧縮データの復元を行うことにより、動画の再生を行うことが可能となる。

図3は音声認識部6の内部構成を示すブロック図である。音声認識部6は、自分の声の登録の必要がなく、誰の声でも認識することができる不特定話者に対応したワードスポットティング機能付き離散単語（太陽、海など単独で用いる単語）の音声認識サブシステムである。この図において、コネクタ62はケーブル2cを介して図2のコネクタ2aに接続される。また、コネクタ63、64にはそれぞれPAD2b、マイク7が接続される。コネクタ62～64は音声認識用LSI61（型名RL5C288：株式会社リコー製）に接続されている。音声認識

用 L S I 6 1 は、マイク 7 から入力される音声信号に基づき音声認識処理を行うとともに、P A D 2 b からのデータを本体側の C P U ブロック 1 0 に伝える。また、P A D 2 b の制御データと音声認識結果とを融合する処理も行う。また、拡張ポート（コネクタ 6 3）に対応ペリフェラルを接続することにより音声とマニュアルの同時操作が可能となる。音声認識用 L S I 6 1 は、音声認識処理のパラメータデータ等を記憶するデータレジスタ 6 1 a、音声認識のための標準パターンが記憶される標準パターンメモリ 6 1 b 及び特定話者の音声認識のための特定パターンメモリ 6 1 c を備える。

音声認識部 6 は、予め各単語ごとに、多数話者のスペクトル系列を分類（クラスタ化）して、各クラスタの中心又は各クラスタに属する音声の平均スペクトル系列を標準パターン（マルチテンプレート）としてもっている。入力音声に対してすべての標準パターンとマッチングを行い、類似度を計算する。類似度とは、認識結果に含まれるデータであって、使用者が発声した音声と認識結果の単語との距離を示す。値が大きいほど類似している。類似度の高い標準パターンの単語を認識結果として出力する。なお、標準パターンの辞書データは、ゲーム機 1 側からダウンロードされる。また、認識結果として単語の類似度とともに、音量取得値も出力される。音量取得値は、入力音声のレベルを示す値であり、これにより使用者が発声した音声の大きさが計測できる。値が大きいほど音量が大きいことを示す。

音声認識部 6 は次のような特徴をもつ。

認識タイプ：離散発声単語、話者タイプ：不特定話者、特微量：メルケプストラム、認識方式：D S T モデル、認識単語数：不特定話者に対して標準 3 0 単語（ただし、総モーラ数によっては認識単語数は変化）、認識処理時間：発声終了

後 0.2 秒以上（アプリケーションにより変更可能）、単語発声長：0.15 ~ 2.00 秒（不要語を含む）、クラスタリング方式：ビットアサイン方式、クラスタ数：最大 16、リジェクトレベル：8 段階（0（低レベル）～7（高レベル））、ワードコネクション：最少 30 通り（ただし、モーラ数によっては組み合わせ数は変化する）、マイクアンプゲイン：可変。

図 4 は、このゲーム機 1 と音声認識部 6 を使ったシステムの構成例である。詳細は後述するが、遊戯者（PLAYER）は TV 受像機 5 の画面を見ながら PAD 2 b を操作したり、マイク 7 に向かってコマンドを発することにより様々な操作を行う。

図 5 は、音声認識の基本処理のフローチャートである。なお、以下の説明において、「単語」とは文法的に厳格な意味での単語を示すのではなく、人の名前や数字等の短いものや、文や文章等の長いものを含めて音声認識の際の一単位を「単語」と標記する。

まず、認識コマンドをデータレジスタ 61a にセットする（ステップ S10）。次に、音声認識用 LSI 61 から認識結果が出力されるまで待つ（ステップ S11）。

認識コマンドを受け取った音声認識用 LSI 61 は、所定レベル以上の音量の音声が一定時間継続すると、認識結果をデータレジスタ 61a に格納するとともに、認識結果が出力されたことをゲーム機本体 10 に知らせる。

認識結果が出力されると、データレジスタ 61a から複数の標準パターンのうちで最もそれらしいとされた認識 1 位の単語と認識得点とを読み出す（ステップ S12）。認識得点とは入力音声と認識 1 位の単語との類似度を得点により現したものである。認識得点が高いほど入力音声と類似していることを現している。

次に、更に詳しいデータが必要か否か判断し（ステップS13）、データが必要でなければステップS10に戻る。

データが必要であれば、ネクストコマンドをデータレジスタ61aにセットする（ステップS14）。ネクストコマンドにより音声認識用LSI61は認識単語発声時の音量データをデータレジスタ61aに格納する。ゲーム機本体10はデータレジスタ61aから音量データを読み出す（ステップS15）。

次に、更に詳しいデータが必要か否か判断し（ステップS16）、データが必要でなければステップS10に戻る。

データが必要であれば、ネクストコマンドをデータレジスタ61aにセットする（ステップS17）。ネクストコマンドにより音声認識用LSI61は認識2位の単語と認識得点をデータレジスタ61aに格納する。ゲーム機本体10はデータレジスタ61aから認識2位の単語と認識得点を読み出す（ステップS18）。

次に、更に詳しいデータが必要か否か判断し（ステップS19）、データが必要でなければステップS10に戻る。

データが必要であれば、ネクストコマンドをデータレジスタ61aにセットする（ステップS20）。ネクストコマンドにより音声認識用LSI61は認識3位の単語と認識得点をデータレジスタ61aに格納する。ゲーム機本体10はデータレジスタ61aから認識3位の単語と認識得点を読み出す（ステップS21）。

このようにして、音声認識用LSI61から認識結果として認識1位の単語と認識得点を得ることができ、必要に応じてネクストコマンドにより更に詳しい認識結果のデータを得ることができる。

以上の音声認識処理により、音声認識部6は次のような機能をもつ。

- ・デフォルト不特定話者音声認識機能

音声認識部 6 は、所定のシーケンスでデフォルト認識モードとノーマル認識モード（後述）を切り替える。デフォルト認識モードは、装置の電源投入直後において表示される、CD-ROMなどを操作するための初期画面を音声により操作することを目的としたモードである。音声認識用LSI 61に内蔵される図示しないROMにPAD 2bの各ボタンに対応した単語辞書を持ち、対応した単語が認識された場合は、各ボタンに対応したビットをアクティブ（0）にする。このモードにより、電源投入直後等において初期設定したり、CD-ROMを用いたゲームスタート等を音声を通じて行うことができる。

単語辞書と対応ボタンの関係を以下に示す。

デフォルト単語辞書と対応ボタンの関係

パッドボタン	単語辞書
Right	なし
Left	なし
Down	なし
Start	「ゲームスタート」
Aボタン	「決定」
Cボタン	「Cボタン」
Bボタン	「取消」
Rボタン	「曲送り」
Xボタン	「くり返し」
Yボタン	「停止」
Zボタン	「再生」
Lボタン	「曲戻し」

- ・ノーマル不特定話者音声認識機能

辞書をゲーム機1のCD-ROMよりロードし、認識を行うゲームでの使用を目的としたモードである。辞書データロードコマンドを発行し、辞書をCD-ROMより音声認識用LSI61内の図示しないRAMにロードし、認識開始コマンドを発行すること（後述のリジェクトレベルと使用クラスタの設定が必要）により認識動作を開始する。認識停止コマンドを発行することにより認識動作を停止させることができる。

この場合において、辞書データのダウンロードは、パワーオン又はリセット時に自動的に行うようにしてもよい。音声認識部6にパワーオンリセットがかからると、CD-ROMからシステム内蔵RAMにデータを自動で転送する。その後、標準パッドデータとの置換を行う。このようにダウンロードを自動的に行うと、外部操作が不要なため、未対応のアプリケーションを使用したときも標準データをアプリケーションで使用することができる。また、標準データを予めシステムに備えて共用することにより各種アプリケーションのデータ量削減及びアプリケーション作成者の負担の軽減が図れる。

- ・ワードスポットティング機能

音声認識部6は、ワードスポットティング機能を有し、「あーパンチ」の「あー」などの不要語を含んだ語から辞書にある単語を分離することができる。

- ・キャッチアップ機能

音声認識部6は、標準パターンメモリ61bに基づき不特定話者音声認識を採用している。しかし、個性の強い発音をするユーザーあるいは特定の単語によっては認識しにくい場合があるかもしれない。そこで、認識しにくい単語をユーザー個別の音声登録に置き換えることにより、認識率の向上を図るようとしてもよ

い。この場合、特定話者音声認識となる。

具体的手順は次のように行う。データレジスタ 61a へ音声データ転送コマンドを発行する。次に、マイク 7 に所望の音声を入力する。そして、入力された音声を特定パターンメモリ 61c に格納する。

・マイクゲイン変更機能

マイク 7 のゲインを変更することができる。データレジスタ 61a へゲイン設定コマンドを発行することにより設定を行う。この機能により、話者の発声音量、周囲の騒音等に応じて適切な認識を行うことができる。

・リアルタイム音量出力機能

ノーマル認識モードには、音声レベルをリアルタイム（20ms 毎）に出力する音スイッチモードと通常の音声認識を行う認識モードの 2 つがある。前者のリアルタイム音量出力モードでは、単に音声レベルを出力するだけなので、非常に高速な処理が可能である。モード音量出力コマンドを発行することにより、8bit (0 ~ 255) の音声レベル信号を出力する。

このモードの用途として、次のようなことが考えられる。

マイク 7 あるいは PAD 2b からのコマンド入力の後に、そのコマンドによる操作の程度を音声レベル信号で決める。例えば、画面上のキャラクタに「前進」を命令した後に、その移動速度を速くするために遊戯者に「はやく、はやく」と叫ばせる。あるいは、画面上のキャラクタの泣き笑い等の表情をつけるとき、どの程度激しい表情にするかを遊戯者の「アーッ」という叫び声の大きさで決めるようとする。あるいは、画面上のキャラクタが危険地帯を歩いているときに、敵キャラクタの突然の出現等の突発的な危険状態が生じたとき、遊戯者が「危ない」とか「キャー」等の叫び声を上げたときに、その大きさに応じて、キャラクタの

動くスピードを遅くしたり、身構えさせたり、姿勢を低くさせたり、気絶させたりする。

また、音声レベルにより程度を変えるだけでなく、動作の種類を変えるようにしてもよい。例えば、画面上のキャラクタに「前進」と命令した後に「行け」と小さな声で命令したときは、キャラクタは歩き出し、一方、「行けー！！」と叫んだときは、キャラクタはジャンプしたり、飛行能力があれば空を飛んだりするようにしてもよい。

要するに、音声レベルにより動作の程度を増加あるいは減少させるように制御したり、あるしきい値を越えたかどうかにより動作の種類を変えるようにすることが考えられる。

・リジェクト機能

音声認識時に音声認識用 L S I 6 1 から類似度とともに出力される得点（レベル）が一定レベル以下であるときに、その認識結果を棄却する機能である。例えば、得点 < (リジェクトレベル + 1) * 8 であるとき、その認識 1 位の類似度の結果は出力されない。リジェクトレベルは、例えば、2 程度である。この機能により、例えば話者以外の人の声やノイズ等により誤って認識することを防止できる。

・クラスタ機能

これは一場面の 30 単語を最大 16 のグループにわけ、任意のグループ（複数選択可能）を選択して認識を行うことにより認識率及び認識速度を向上させる機能である。

例えば、ロールプレイングゲーム（RPG）の戦闘場面でよく用いられる呪文等を認識するために、呪文それぞれの種類ごとに攻撃魔法（クラスタ 1）、回

復魔法（クラスタ2）、直接攻撃（クラスタ3）、その他（クラスタ4）、職種（クラスタ5）というクラスタを作る。そして、戦闘場面の経過ごとに順番に適宜クラスタを選択していく。

戦闘を開始する前は、職種（クラスタ5）の中の言葉のいずれかであると認識する。認識結果が「戦士」であれば、次は当然、武器をもって直接攻撃することになるから、直接攻撃（クラスタ3）から認識する。他方、職種の認識結果が「魔法使い」であるときは、攻撃魔法（クラスタ1）から認識する。

このように、前回の認識結果に基づき、場面毎に関連の深い適当なクラスタを選択し、音声信号がこの選択されたクラスタの言葉のいずれかであると判定する。この方法によれば、まったく無関係の言葉であると誤認するおそれが少くなり、認識の信頼性が高くなるとともに、類似度を求める認識処理を行う対象がその選択されたクラスタ内に限定されるので処理演算回数が減少し、したがって処理速度の点でも有利である。

・ワードコネクション機能

連続で発声した2単語を1組として認識の結果を出力する機能である。1単語目の類似度と2単語目の類似度の合計が高いものを出力する。この場合、予め2単語の結合データを辞書に入れておく必要がある。

例えば、辞書に、猿、攻撃、雉、防御、犬、魔法、鬼、逃げる、桃太郎の語彙が存在し、さらに、結合データとして、猿-攻撃、猿-防御、猿-逃げる、雉-攻撃、雉-防御、雉-逃げる、犬-攻撃、犬-防御、犬-逃げる、が格納されているとする。このように、辞書内において、単語とこれらの結合データが格納されている。

この場合において、ユーザーが「猿、攻撃」（間を若干開けること）と発声し

たとすれば、この時の音声認識部6の動作は（9語から3語を選んでの順列＊2通りのうちの）例えば次のケースのようになる。

ケース 1

1 単語目類似度	2 単語目類似度	認識結果
1位 猿	1位 攻撃	
2位 雉	2位 魔法	雉 - 攻撃
3位 鬼	3位 逃げる	

ケース 2

1 単語目類似度	2 単語目類似度	認識結果
1位 雉	1位 攻撃	
2位 猿	2位 魔法	雉 - 攻撃
3位 鬼	3位 逃げる	

ケース 3

1 単語目類似度	2 単語目類似度	認識結果
1位 鬼	1位 攻撃	
2位 猿	2位 魔法	猿 - 攻撃
3位 犬	3位 逃げる	

ケース 4

1 単語目類似度	2 単語目類似度	認識結果
1位 鬼	1位 魔法	
2位 猿	2位 攻撃	猿 - 攻撃
3位 犬	3位 逃げる	

以上のワードコネクション機能使用時、音声認識部6は類似度1～3位の候補

の中から出力を選択する。1単語目類似度1、2、3位の単語と2単語目類似度1、2、3位の組み合わせで合計の類似度が大きいものが選択される。しかし、ケース3、4のように類似度1位に結合データを持たない単語がきたときは、これを棄却して結合データを優先して出力する。

次に、音声認識部を用いたゲーム機の動作について説明する。

この種のゲーム機は、通常、パッド2bのボタンにより操作されるが、これに代えてあるいはこれと併せて音声認識装置を用いることにより新しい使用方法が可能になる。音量の差の要因を識別できれば、これと音声認識とを組み合わせることにより、新しいプレイが可能になる。

普通の音声認識だと発音された音声が認識データと一致したか否かの判断を行い、その結果を出力するのみであるが、これにさらに発音された音声がある基準より大きかったか、それとも小さかったかの判断結果を付加することにより応用範囲を広げることができる。この具体例については後述する。

発生された音は、マイクロフォン7を通して音声及び音量を認識する音声認識部6に送られ、ここで認識データの比較及び音量基準値の比較を行い、その結果をゲーム機本体1に送る。ゲーム機本体1ではその結果に従つていろいろな動作を行う。なお、ゲーム機本体1から音声認識部6に対して、音量基準値のセットができる。また、音量基準値は、調整可能なデフォルト値を持っている。

次に、音声認識と音量値との組み合わせによる動作の具体例について、ゲーム機の場合を例にとり説明する。

<音声レベルにより動作の程度を制御>

例えば、画面上のキャラクタに「前進」を命令した後に、その移動速度を速くするために遊戯者に「はやく、はやく」と呼ばせる。あるいは、画面上のキャラ

クタの泣き笑い等の表情をつけるとき、どの程度激しい表情にするかを遊戯者の「アーッ」という叫び声の大きさで決めるようとする。あるいは、画面上のキャラクタが危険地帯を歩いているときに、敵キャラクタの突然の出現等の突発的な危険状態が生じたとき、遊戯者が「危ない」とか「キャー」等の叫び声を上げたときに、その大きさに応じて、キャラクタの動くスピードを遅くしたり、身構えさせたり、姿勢を低くさせたり、気絶させたりする。あるいは音量により強いキックを出すか、それとも弱いキックを出すか選択する。

ところで、音量レベルの意味として互いに異なるいくつかのとらえ方が可能である。遊戯者の声の出し方は様々であるが、概念的にとらえると図6及び図7のような例が考えられる。これら図の縦軸は音量レベルであり、横軸は時間である。図6において、最初の音量が最も大きく、徐々に下がっていっている。「キャー」という悲鳴のような興奮状態の声はこのように変化すると考えられる。これに対し、図7において、徐々に音量が大きくなつてやがてピークを迎え、その後、徐々に音量が小さくなっている。通常の音声はこのように変化すると考えられる。

音量レベルを次のように定義する。

(1) 平均音量

図6及び図7の信号全体についての平均を求め、これを(平均)音量とする。具体的には、マイクの入力信号に対して積分回路を適用すればよい。平均音量は、声が通常の状態であるか興奮状態であるかにかかわらず声の全体(音量×持続時間)の値を示すので、キャラクタの動くスピードのように、割と単純な動作の程度を制御する用途に適する。

(2) ピーク音量

信号のピーク値(図6及び図7の点P)のレベル Δy を求め、これをピーク音

量とする。具体的には、サンプルホールド回路を用いて信号をサンプルしつつ、最大値を保持し続ければよい。ピーク音量は最大の音量を示すから、複数の遊戯者間で声の大きさを競うような用途に適する。

(3) 立ち上がり時間

信号のピークまでの時間（図6及び図7の Δt ）を求め、これを立ち上がり時間とする。上述のピークホールド回路のホールド時刻を知ることにより簡単に求めることができる。あるいは、立ち上がりの角度 θ 、 $\Delta y / \Delta t$ で立ち上がり時間を定義するようにしてもよい。これは微分回路により実現できる。立ち上がり時間は、いわば声の緊迫度に対応すると考えられるから、例えば、立ち上がりが急なときにはキャラクタに身構えさせたり退避行動をさせたりし、立ち上がりが急でないときにはキャラクタを速く歩かせるようにすることが考えられる。

(4) 持続時間

信号の持続時間（図6及び図7のT）を求め、これを持続時間とする。具体的には、図示しないタイマにより、音声信号のうちで予め定められたしきい値以上のレベルの継続時間を計測する。持続時間は、単純に声の長さを競うような場合に適用できる。

<音声レベルにより動作の種類を制御>

また、音声レベルにより程度を変えるだけでなく、動作の種類を変えるようにしてもよい。例えば、画面上のキャラクタに「前進」と命令した後に「行け」と小さな声で命令したときは、キャラクタは歩き出し、一方、「行けー！！」と叫んだときは、キャラクタはジャンプしたり、飛行能力があれば空を飛んだりするようにもよい。

要するに、音声レベルにより動作の程度を増加あるいは減少させるように制御

したり、あるしきい値を越えたかどうかにより動作の種類を変えるようにすることが考えられる。

<音声の種類により動作の種類を制御>

子供の声と大人の声、あるいは男性の声と女性の声とを区別して、この結果に基づき操作を変えるようにしてもよい。例えば、子供のキャラクタと大人のキャラクタが混在しているとき、声の種類の応じて子供あるいは大人のキャラクタのみを操作する。男性と女性の場合も同様である。このように制御すれば、複数の遊戯者が同時に声を出したときでもプレイ可能となり、ゲームに幅が出てくる。

ところで、子供の声と大人の声とは、その周波数分布の点で異なるから識別可能である。例えば、フーリエ変換して周波数分布を求めるにより、あるいは簡単に複数のフィルタの出力を比較することにより識別することができる。男性の声と女性の声も周波数分布の点で異なるから同様の手段により識別できる。

<その他>

上述の音声レベル又は／及び音声の種類と、操作パッド2 bのボタン操作との組み合わせにより特有の動作をさせるようにしてもよい。例えば、大きな声+Aボタン操作により、そのときの同じ技が繰り返すように制御することが考えられる。また、複数のボタンを同時に押した場合にある必殺技が出せるような格闘ゲームのような場合、あるボタンには音声入力を、他のボタンには通常のパッド入力を割り当れば、音声とパッド入力の組み合わせによって、パッドのみを使用したキー入力に比較して、より簡単な入力で同様なキー入力が行える。

なお、上記の音声レベルを用いる場合において、前述のリアルタイム音量出力機能を用いてもよい。すなわち、音声認識モードとリアルタイム音量出力モードを組み合わせる。まず音声認識を行って対応する動作をさせた後、リアルタイム

音量出力モードに入り、音声レベルのみを出力する。このようにするとレスポンスが非常に良くなる。

以上のように、この発明の実施の形態によれば、小さく発せられた言葉に対しては少ない動き、大きく発せられた言葉に対しては大きな動きというように、発声音量に基づき動きの程度を制御することができる。また、発声音量に基づき動作の種類を制御することができる。また、音声の種類に基づき動作の種類を制御することができる。

形態 2.

上記発明の実施の形態 1 の装置は、従来の制御装置であるパッドに加えて、音声認識装置を組み合わせることにより操作に奥行きを持たせたものである。これに対して、パッドを音声認識装置で置き換えるようにしてもよい。

図 3 の音声認識部 6 は、前述のように、デフォルト不特定話者音声認識機能及びノーマル不特定話者音声認識機能を持っている。したがって、電源を投入した直後のデフォルト状態でも所定の音声とパッドのボタンとは対応するから、音声認識に対応しないソフトであっても音声認識装置を利用することができる。また、ノーマル不特定話者音声認識機能の辞書データロードコマンドを発行し辞書を C D - R O M からダウンロードすることにより、パッドのボタンと音声認識結果との関係について、そのソフト特有の対応をとることができる。

次に、ノーマル特定話者音声認識機能を用いたときの処理動作について説明する。

まず、音声認識部 6 は辞書データロードコマンドを発行することにより、C D - R O M から辞書データをダウンロードする。

次に、辞書データに基づきパッド 2 b のボタンと所定の音声の対応関係を定める。まず、最初に、本体のキースイッチに対応させる音声を音声データ RAM に定義する。例えば、「すすめ」という音声が入力されると、「上」を示すキーデータを出力するように定義される。このような状態で、図 1 の構成のマイクロフォンより「すすめ」の音声が入力されると、そのデータをデジタルデータとして音声比較部に取り込んだ後に、音声データ RAM の中に同じデータが存在するかどうかを判定する。もし、マイクロフォンより入力された音声と、音声データ RAM の中に同じデータが存在すれば、そのデータにしたがって「上」を示すキーデータがゲーム機本体 1 へ送信される。

音声信号が入力されると、音声認識部 6 は音声を認識し、上記の対応関係に基づき、通常のゲームパッド操作に対応するキーデータをゲーム機本体 1 に出力する。例えば、「ゲームスタート」と言えば、スタートボタンが押されたときと同じ信号を出力する。なお、パッド上の各ボタンに対して遊技者が好きな音声を登録することも可能である。例えば、そのための専用のコマンドを入力し、パッドデータそれぞれに所望の単語を選択あるいは入力して登録すればよい。このとき、TV 画面上で文字を選択すれば不特定話者音声認識が可能になるし、マイク 7 から所望の単語を入力すれば特定話者認識が可能になる（キャッチアップ機能）。

今までのゲーム機へのデータ入力装置では、手や指等を使用してゲームパッドに取り付けられたスイッチを押すことによって、ゲーム機本体へのデータの入力を行っていたが、以上のようにすれば、音声による入力でも同様にゲームをコントロールできる。

また、音声認識部 6 を利用すれば、ゲームパッドを用いた場合とは異なる遊び方ができる。通常、格闘ゲームなどで難度の高い技を出すとき、高度なパッド操

作を要求されるが、音声認識部6を用いた場合、高度な音声入力（早口とか難しい発音など）をしないと技が出ないようにする。また、前述の音声レベルや音声種類の識別を組み合わせるようにしててもよい。

このように、音声を使用してキー入力をを行うことで、よりゲームに親しみがもてるようになる。例えば、ボクシングゲーム等でAボタンが左パンチ、Bボタンが右パンチというキーアサインになっているとき、A、Bボタンを使用してパンチを出すよりも、「左、右」という音声でパンチを出す方がゲームが白熱する。

発明の実施の形態3.

音声認識部6は、音声認識後に認識単語番号と類似度を出力する。ところで、認識開始コマンド発行時にリジェクトレベルが設定されている場合、認識結果の類似度が設定値より低いときは結果を出力しないことが可能である。この機能により、例えば話者以外の人の声やノイズ等により誤って認識することを防止できる。

誤認識を防止することの他にも、リジェクトレベルには次のような応用方法がある。

- ・場面の状況によりリジェクトレベルを変える

例えば、アドベンチャーゲームにおいて、なんらかの誤認識が起きるとプレイヤーが死んでしまい、ゲームオーバーになってしまうという場面では、リジェクトレベルを高く設定する。このようにすれば、レスポンス速度を犠牲にしても出力結果の確実性が向上する。

他方、ゲームの進行にさして関らない部分はリジェクトレベルを下げることにより、出力結果の確実性は低下するものの、レスポンスを良くすることができる。このような設定の仕方をゲーム開発者が任意に行えるので、パッドを用いる場

合と比べてさまざまな制御が可能になる。

発明の実施の形態 4.

音声認識部 6 は、音声認識後に認識単語番号と類似度を出力する。したがって、類似度を用いて動作の程度や動作の種類を制御するようにしてもよい。

例えば、ある単語が認識されたとき、音声認識部 6 が出力する類似度の大きさに基づき、レスポンスに差をつける。例えば、類似度 80 点のパンチの威力は強力にし、一方、10 点のパンチの威力は弱くする。このように、類似度によりパンチの威力が違うようになる。すると、不特定話者音声認識を行う場合、遊戯者の発音によりキャラクタの動作、ゲームの進行が異なるようになり、遊戯者は標準データの音声に合わせた方が有利となる。このような観点（誰が一番標準的な発音ができるか、音声認識に適した発音ができるか等）からゲームを楽しむこともできる。したがって、ゲームの楽しみ方に深みがでる。

なお、発明の実施の形態 1 で説明した、音声レベルの関係において説明した具体例について、この実施の形態 4 の類似度による制御の考え方を適用することもできる。

発明の実施の形態 5.

以上の実施の形態において音声認識部を中心に説明したが、この音声認識部を用いて行うゲームについて説明する。

(1) 動物調教アシストアクションゲーム

音声認識を使用し、犬、ネコ、馬などの動物達とのコミュニケーションを楽しむゲームである。

一つのキャラクターだけではなく、複数のキャラクターを直接的または間接的に操作できるという音声認識ならではのゲームである。この種のゲームをボタン

操作で行うと味気なく、実際と同様に「声をかける」という操作が適している。

また、実際の調教において動物がなかなか言うことを聞かないことがある。音声認識がうまくいかないことがあっても、このときのもどかしさ自体もゲーム性に含まれるので、面白さが却って増加する。呼び掛けたときのキャラクターのリアクションの多彩さが面白い。

あるいは、プレーヤーがお供にしている猿、雉、犬等の動物達が勝手に動きまわるアクションゲームも考えられる。桃太郎物語のように、プレーヤーはフィールド上にいる動物達に実際に話し掛けることで仲間にすることができる。たとえば、動物を仲間にするには仲間にしたい動物がプレーヤーのキャラクタを見ているときに何でも良いから呼び掛ける。例えば、犬だったら「ぼち！　ぼち！　ぼち！」というように何度も呼び掛けているとその声に反応するようになる。そこで、餌をやるとその動物はプレーヤーの仲間になってくれる。

プレーヤーが歩いていくと動物は後からついてくるが、なにか興味のあるものを見つけると立ち止まったり、そっちの方へ行ってしまおうとする。そういう場合には、「ぼち！　早く来い！」などの呼び掛けをすることによって引き戻す。

動物達は常にいろいろな動きをしている。プレーヤーが動物のある動作に対して何回か同じ言葉をかけると動物はその言葉に対してはある動作を必ずするようになる。この積み重ねでプレーヤーはだんだんと動物を操れるようになっていく。あるいは、トラップ、エネミーをクリアしていくアクションゲームも考えられる。羊飼いが羊達を追い立てるようなゲームである。

このゲームでは様々なトラップが用意されているが、プレーヤーは動物を操ることによって、トラップをクリアしていく。プレーヤーが簡単に気付くようなトラップでも動物はひっかかるってしまうのでそれを教える。仲間の動物が増えるほど

ど難しい。トラップをクリアして先へ進むために、どの動物をどのようにして使うかをプレーヤーに考えさせる。

あるいは、熱帯魚飼育ゲーム、錦鯉飼育ゲーム等である。声だけに限らず手を叩く等の動作を通じて魚達を集め。手の叩き方を工夫することにより、魚が様々な動作をするようにプログラムしておけば面白い楽しみ方が可能になる。

(2) 多人数早答えクイズ

ゲーム機で行うクイズ番組である。複数の参加者がいて、ゲーム機が順番に問題を出していって参加者間で得点を競うゲームである。この場合、答えの語彙が(1) (2) . . . とか (ア) (イ) . . . など限定されるので、処理速度の点で有利である。

(3) 子供の教育ゲーム

子供のキャラクタに言葉を覚えさせたり、躾をしたりする。この種のゲームはコミュニケーションが目的であり、音声を使うとより面白味が増す。

(4) フライトシミュレーションゲーム

ウエポン等の装備の選択、サブパイロットへの命令に音声認識部を用いればより実感的である。

(5) 麻雀ゲーム

ティー、ロン、パービン等を発声して手を選ぶ。

(6) その他

将棋ゲーム（投了、2六銀等）、テーブルゲーム、パーティゲーム（マウスの代わり、対話式ゲーム）、ナンバーゲーム（彼女、お茶しない、OK等）、ドライブゲーム（無線のやり取り）、RPG（ロールプレイングゲーム）のショートカット、対話式仮想家族ゲーム（老人用）、戦国シミュレーション等が考えられ

る。

尚、記憶媒体とは、何等かの物理的手段により情報（主にデジタルデータ、プログラム）が記録されている、既述のCD-ROM、カートリッジ等であって、コンピュータ、専用プロセッサ等の処理装置に所定の機能を行わせることができるものである。要するに、何等かの手段でもってコンピュータにプログラムをダウンロードし、所定の機能を実行させるものであればよい。

例えば、フロッピーディスク、ハードディスク、磁気テープ、光磁気ディスク、CD-ROM、DVD、ROMカートリッジ、バッテリバックアップ付きのRAMメモリカートリッジ、フラッシュメモリカートリッジ、不揮発性RAMカートリッジ等を含む。

また、電話回線等の有線通信媒体、マイクロ波回線等の無線通信媒体等の通信媒体、インターネット等の通信媒体記憶媒体に含まれる。

産業上の利用可能性

以上のように、この発明によれば、ゲーム機の周辺装置として用いられる音声認識装置であって、音声入力手段と、前記音声入力手段の音声信号出力と予め定められた音声認識用辞書のデータとを比較することにより遊戯者の音声を認識するとともに、認識結果に基づきゲームに関する制御信号を生成する音声認識部とを備えるので、ゲームの遊戯者は音声によるプレイが可能になる。

また、この発明によれば、前記音声認識部は、前記音声認識用辞書として、不特定話者用に予め設定された不特定話者音声認識用辞書及び遊戯者により設定される特定話者音声認識用辞書を備え、動作の初期状態において前記不特定話者音声認識用辞書を選択するので、特定話者に対応できないゲームソフトについても

不特定話者音声認識により音声によるプレイが可能になる。専用ソフトに限らず多くのソフトに対応できて、この装置の適用範囲が広がる。

また、この発明によれば、前記音声認識部は、複数の遊戯者それぞれに対応する複数の特定話者音声認識用辞書を備え、遊戯者の操作によりこれら特定話者音声認識用辞書から認識処理に使用する辞書を選択するので、適当な辞書により特定話者音声認識ができて、音声認識の精度が向上する。

また、この発明によれば、前記音声認識部に接続されるゲーム機用の制御部を備え、前記音声認識部は、音声の認識結果と前記制御部の制御信号とを組み合わせてゲームに関する制御信号を生成するので、通常のパッド等と音声認識とを組み合わせて様々な操作が可能となり、面白さに深みができる。

また、この発明によれば、前記制御部は、通常の動作をさせる制御信号を出力し、前記音声認識部は、特殊な動作をさせる制御信号を生成するので、通常のパッド等では難しかった特殊な動作を実行させやすくなる。

また、この発明によれば、前記音声認識部は、前記音声入力手段の音声信号出力と前記音声認識用辞書の内容との類似の程度を示す類似度及び前記音声信号の状態を示す値を出力するとともに、前記音声認識部は、ゲームにおける動作の種類を選択するために前記類似度に基づき音声認識を行う第1の動作モードと、動作の状態を設定するために音声認識を行わずに前記音声信号の状態を測定する第2の動作モードとを備え、第2の動作モードの処理を高速に行うことができるので、応答が迅速になりゲーム操作において有利である。

また、この発明によれば、前記音声認識部は、前記音声入力手段の音声信号出力と前記音声認識用辞書の内容との類似の程度を示す類似度及びその音量レベルを出力し、この音量レベルを予め定められたリジェクトレベルに基づき評価し、

評価結果に基づき前記音声認識部の認識結果をリジェクトするので、適切に認識できない場合の認識を避けて誤った認識を減らすことができる。

また、この発明によれば、前記リジェクトレベルをゲームの種類あるいはゲームステージごとに設定するので、場面に応じて適当なリジェクトレベルを設定できてゲームプレイ上、便利である。

また、この発明によれば、前記音声認識部は、前記音声入力手段の音声信号出力と前記音声認識用辞書の内容との類似の程度を示す類似度に基づき認識とともに、認識結果に基づき生成された制御信号によるゲームにおける動作の状態を前記類似度に応じて変化させて、遊戯者の発音によりキャラクタの動作、ゲームの進行が異なるようになり、ゲームの楽しみ方に深みができる。

請求の範囲

1. ゲーム機の周辺装置として用いられる音声認識装置であつて、音声入力手段と、前記音声入力手段の音声信号出力と予め定められた音声認識用辞書のデータとを比較することにより遊戯者の音声を認識するとともに、認識結果に基づきゲームに関する制御信号を生成する音声認識部とを備える音声認識装置。
2. 前記音声認識部は、前記音声認識用辞書として、不特定話者用に予め設定された不特定話者音声認識用辞書及び遊戸により設定される特定話者音声認識用辞書を備え、動作の初期状態において前記不特定話者音声認識用辞書を選択することを特徴とする請求項1記載の音声認識装置。
3. 前記音声認識部は、複数の遊戸それぞれに対応する複数の特定話者音声認識用辞書を備え、遊戸の操作によりこれら特定話者音声認識用辞書から認識処理に使用する辞書を選択することを特徴とする請求項2記載の音声認識装置。
4. 前記音声認識部に接続されるゲーム機用の制御部を備え、前記音声認識部は、音声の認識結果と前記制御部の制御信号とを組み合わせてゲームに関する制御信号を生成することを特徴とする請求項1乃至請求項3いずれかに記載の音声認識装置。
5. 前記制御部は、通常の動作をさせる制御信号を出力し、前記音声認識部は、特殊な動作をさせる制御信号を生成することを特徴とする請求項3記載の音声認識装置。

6. 前記音声認識部は、前記音声入力手段の音声信号出力と前記音声認識用辞書の内容との類似の程度を示す類似度及び前記音声信号の状態を示す値を出力するとともに、前記音声認識部は、ゲームにおける動作の種類を選択するために前記類似度に基づき音声認識を行う第1の動作モードと、動作の状態を設定するために音声認識を行わずに前記音声信号の状態を測定する第2の動作モードとを備えることを特徴とする請求項1乃至請求項5いずれかに記載の音声認識装置。

7. 前記第2のモードにおいて、前記音声信号の状態を、信号の一部あるいは全体の平均音量レベルとし、この平均音量レベルに基づき動作の状態を設定することを特徴とする請求項6記載の音声認識装置。

8. 前記第2のモードにおいて、前記音声信号の状態を、信号のピーク音量レベルとし、このピーク音量レベルに基づき動作の状態を設定することを特徴とする請求項6記載の音声認識装置。

9. 前記第2のモードにおいて、前記音声信号の状態を、信号の立ち上がり時間とし、この立ち上がり時間に基づき動作の状態を設定することを特徴とする請求項6記載の音声認識装置。

10. 前記第2のモードにおいて、前記音声信号の状態を、信号の持続時間とし、この持続時間に基づき動作の状態を設定することを特徴とする請求項6記載の音声認識装置。

1 1. 前記第2のモードにおいて、前記音声信号の状態を、音声の種類とし、この音声の種類に基づき動作の状態を設定することを特徴とする請求項6記載の音声認識装置。

1 2. 前記音声認識部は、前記音声入力手段の音声信号出力と前記音声認識用辞書の内容との類似の程度を示す類似度及びその音量レベルを出力し、この音量レベルを予め定められたリジェクトレベルに基づき評価し、評価結果に基づき前記音声認識部の認識結果をリジェクトすることを特徴とする請求項1乃至請求項5いずれかに記載の音声認識装置。

1 3. 前記リジェクトレベルをゲームの種類あるいはゲームステージごとに設定することを特徴とする請求項1 2記載の音声認識装置。

1 4. 前記音声認識部は、前記音声入力手段の音声信号出力と前記音声認識用辞書の内容との類似の程度を示す類似度に基づき認識するとともに、認識結果に基づき生成された制御信号によるゲームにおける動作の状態を前記類似度に応じて変化させることを特徴とする請求項1乃至請求項5いずれかに記載の音声認識装置。

1 5. 制御用周辺機器として、請求項1乃至請求項1 4いずれかに記載の音声認識装置を備えるゲーム機。

1 6. 音声信号を受ける第1のステップと、前記音声信号と予め定められた音声認識用辞書のデータとを比較することにより遊戯者の音声を認識する第2のステップと、前記第2のステップの認識結果に基づきゲームに関する制御信号を生成する第3のステップとを備える音声認識方法。

1 7. ゲーム機用の制御部から制御信号を受ける第4のステップを備え、前記第3のステップにおいて、前記第2のステップによる音声の認識結果と前記第4のステップによる制御信号とを組み合わせてゲームに関する制御信号を生成することを特徴とする請求項1 6記載の音声認識方法。

1 8. 前記音声認識部の音声認識動作を実現するプログラムが記憶された媒体。

補正書の請求の範囲

[1997年12月5日(05.12.97)国際事務局受理：出願当初の請求の範囲2, 4, 6, 12, 14及び15は補正された；出願当初の請求の範囲1は取り下げられた；他の請求の範囲は変更なし。(4頁)]

2. (補正後) ゲーム機の周辺装置として用いられる音声認識装置であって、音
5 声入力手段と、前記音声入力手段の音声信号出力と予め定められた音声認識用辞
書のデータとを比較することにより遊戯者の音声を認識するとともに、認識結果
に基づきゲームに関する制御信号を生成する音声認識部とを備え、

前記音声認識部は、前記音声認識用辞書として、不特定話者用に予め設定され
た不特定話者音声認識用辞書及び遊戯者により設定される特定話者音声認識用辞
10 書を備え、動作の初期状態において前記不特定話者音声認識用辞書を選択するこ
とを特徴とする音声認識装置。

3. 前記音声認識部は、複数の遊戯者それぞれに対応する複数の特定話者音声認
識用辞書を備え、遊戯者の操作によりこれら特定話者音声認識用辞書から認識処
15 理に使用する辞書を選択することを特徴とする請求項2記載の音声認識装置。

4. (補正後) 前記音声認識部に接続されるゲーム機用の制御部を備え、前記音
声認識部は、音声の認識結果と前記制御部の制御信号とを組み合わせてゲームに
関する制御信号を生成することを特徴とする請求項2又は請求項3に記載の音声
20 認識装置。

5. 前記制御部は、通常の動作をさせる制御信号を出力し、前記音声認識部は、
特殊な動作をさせる制御信号を生成することを特徴とする請求項3記載の音声認

識装置。

6. (補正後) 前記音声認識部は、前記音声入力手段の音声信号出力と前記音声認識用辞書の内容との類似の程度を示す類似度及び前記音声信号の状態を示す値を出力するとともに、前記音声認識部は、ゲームにおける動作の種類を選択するために前記類似度に基づき音声認識を行う第1の動作モードと、動作の状態を設定するために音声認識を行わずに前記音声信号の状態を測定する第2の動作モードとを備えることを特徴とする請求項2乃至請求項5いずれかに記載の音声認識装置。

10

7. 前記第2のモードにおいて、前記音声信号の状態を、信号の一部あるいは全体の平均音量レベルとし、この平均音量レベルに基づき動作の状態を設定することを特徴とする請求項6記載の音声認識装置。

15

8. 前記第2のモードにおいて、前記音声信号の状態を、信号のピーク音量レベルとし、このピーク音量レベルに基づき動作の状態を設定することを特徴とする請求項6記載の音声認識装置。

20

9. 前記第2のモードにおいて、前記音声信号の状態を、信号の立ち上がり時間とし、この立ち上がり時間に基づき動作の状態を設定することを特徴とする請求項6記載の音声認識装置。

10. 前記第2のモードにおいて、前記音声信号の状態を、信号の持続時間とし、

この持続時間に基づき動作の状態を設定することを特徴とする請求項 6 記載の音声認識装置。

11. 前記第 2 のモードにおいて、前記音声信号の状態を、音声の種類とし、この音声の種類に基づき動作の状態を設定することを特徴とする請求項 6 記載の音声認識装置。

12. (補正後) 前記音声認識部は、前記音声入力手段の音声信号出力と前記音声認識用辞書の内容との類似の程度を示す類似度及びその音量レベルを出力し、この音量レベルを予め定められたリジェクトレベルに基づき評価し、評価結果に基づき前記音声認識部の認識結果をリジェクトすることを特徴とする請求項 2 乃至請求項 5 いずれかに記載の音声認識装置。

13. 前記リジェクトレベルをゲームの種類あるいはゲームステージごとに設定することを特徴とする請求項 12 記載の音声認識装置。

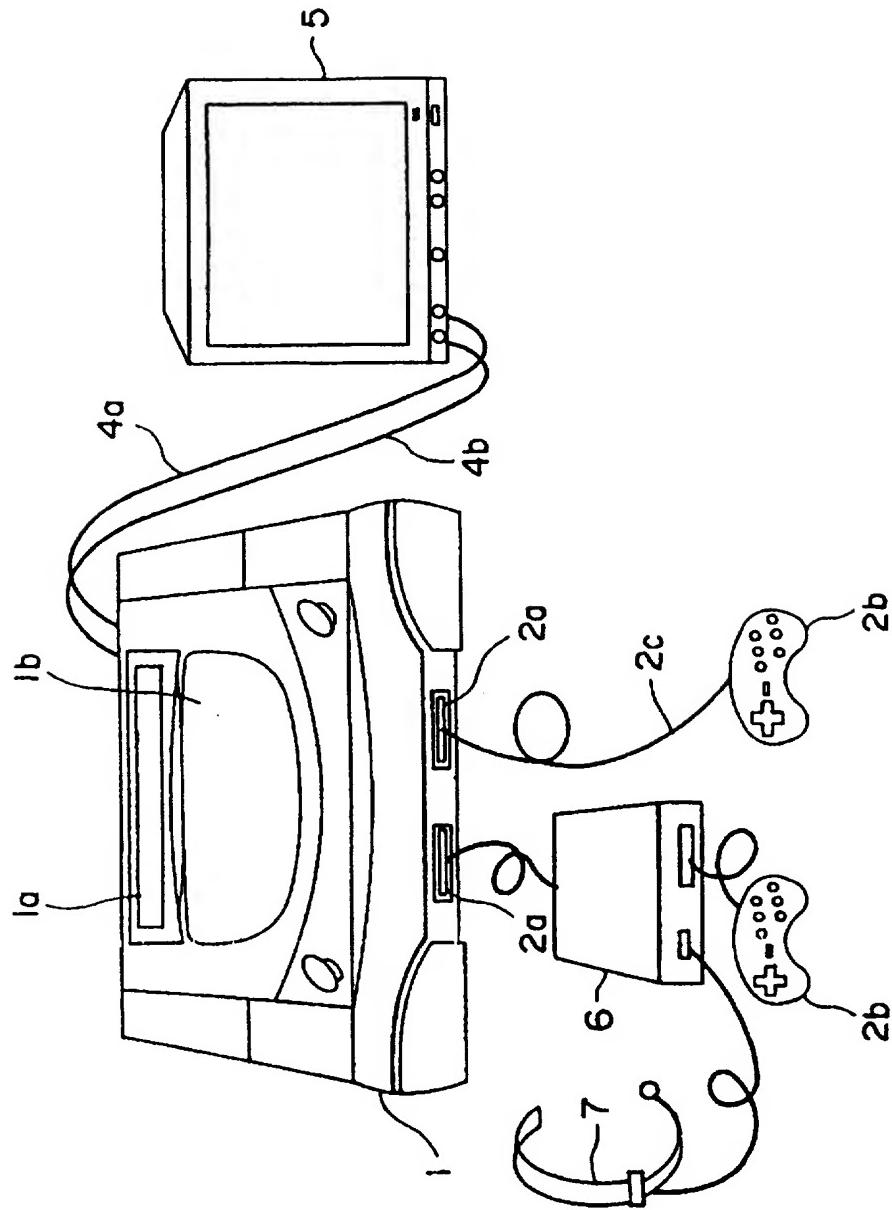
14. (補正後) 前記音声認識部は、前記音声入力手段の音声信号出力と前記音声認識用辞書の内容との類似の程度を示す類似度に基づき認識するとともに、認識結果に基づき生成された制御信号によるゲームにおける動作の状態を前記類似度に応じて変化させることを特徴とする請求項 2 乃至請求項 5 いずれかに記載の音声認識装置。

15. (補正後) 制御用周辺機器として、請求項 2 乃至請求項 14 いずれかに記

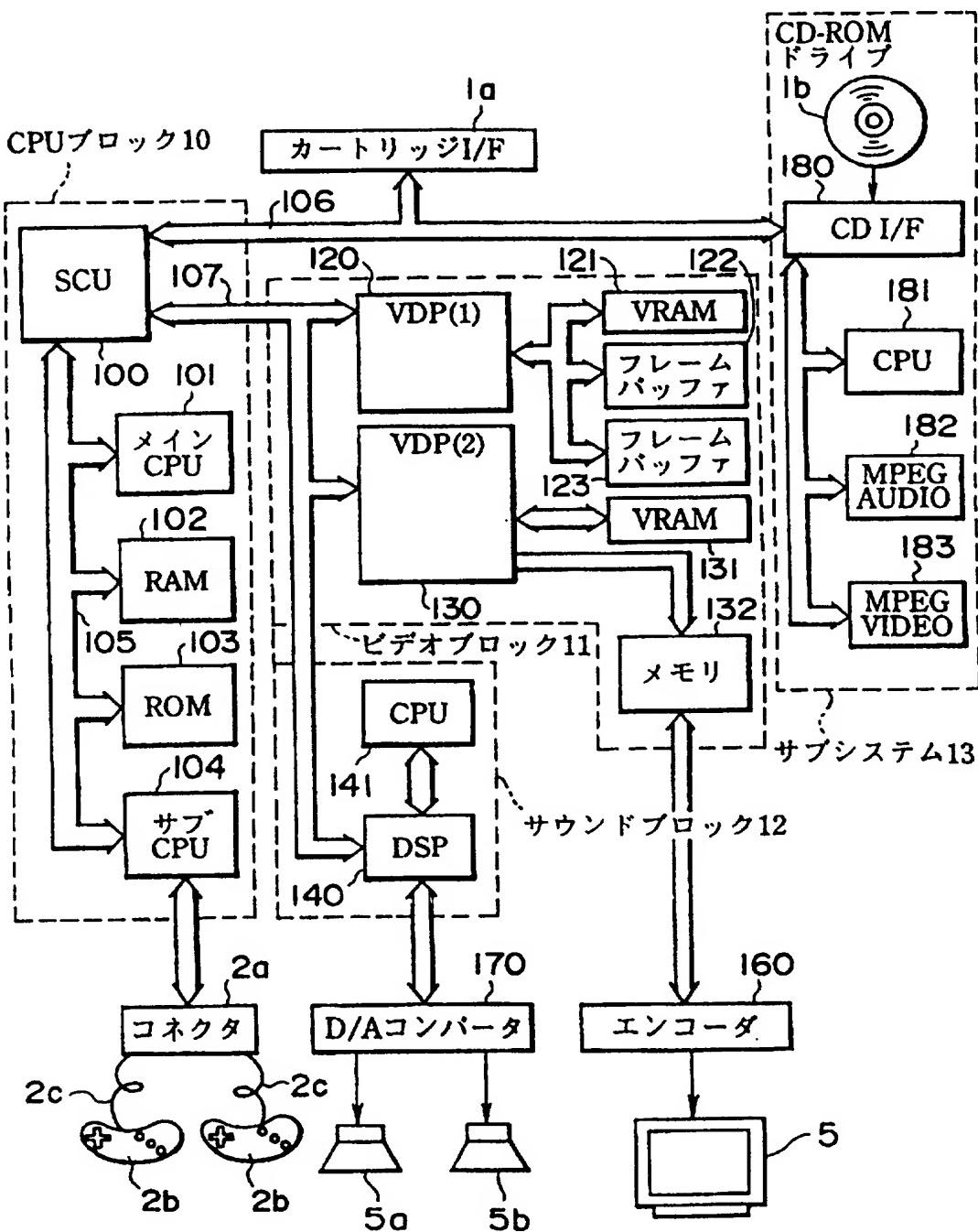
載の音声認識装置を備えるゲーム機。

16. 音声信号を受ける第1のステップと、前記音声信号と予め定められた音声認識用辞書のデータとを比較することにより遊戯者の音声を認識する第2のステップと、前記第2のステップの認識結果に基づきゲームに関する制御信号を生成する第3のステップとを備える音声認識方法。
5
17. ゲーム機用の制御部から制御信号を受ける第4のステップを備え、前記第3のステップにおいて、前記第2のステップによる音声の認識結果と前記第4のステップによる制御信号とを組み合わせてゲームに関する制御信号を生成することを特徴とする請求項16記載の音声認識方法。
10
18. 前記音声認識部の音声認識動作を実現するプログラムが記憶された媒体。

【図 1】

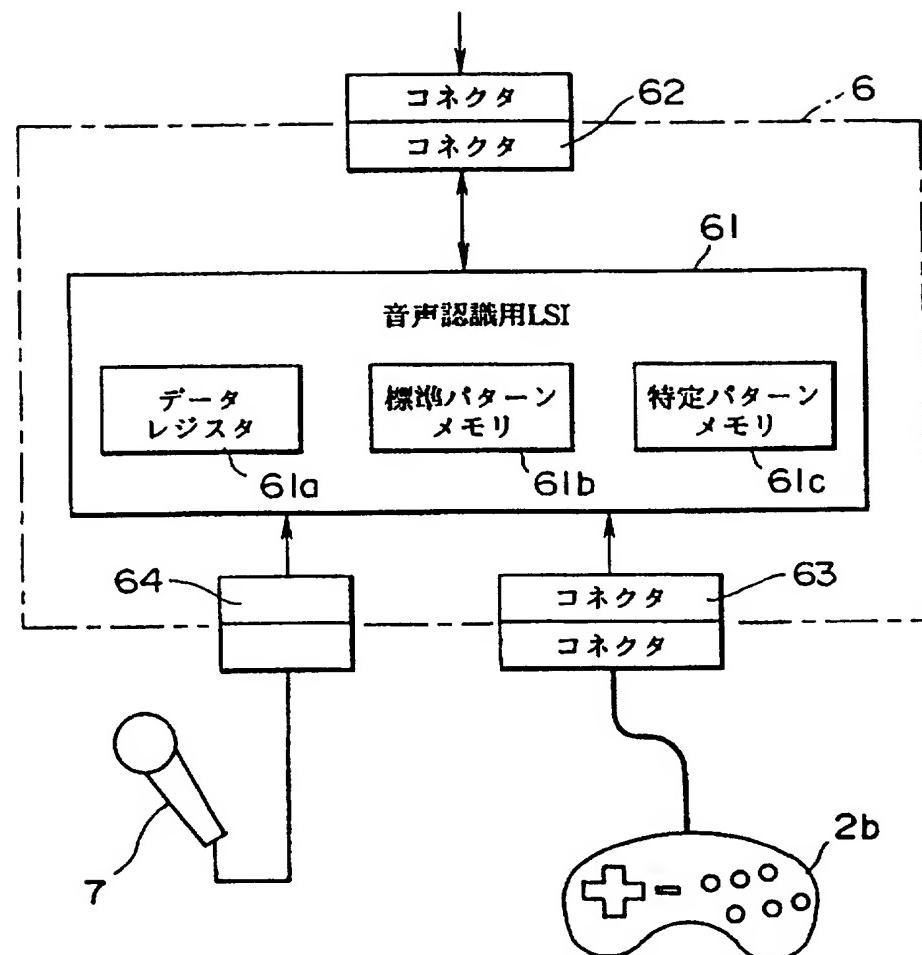


【図 2】

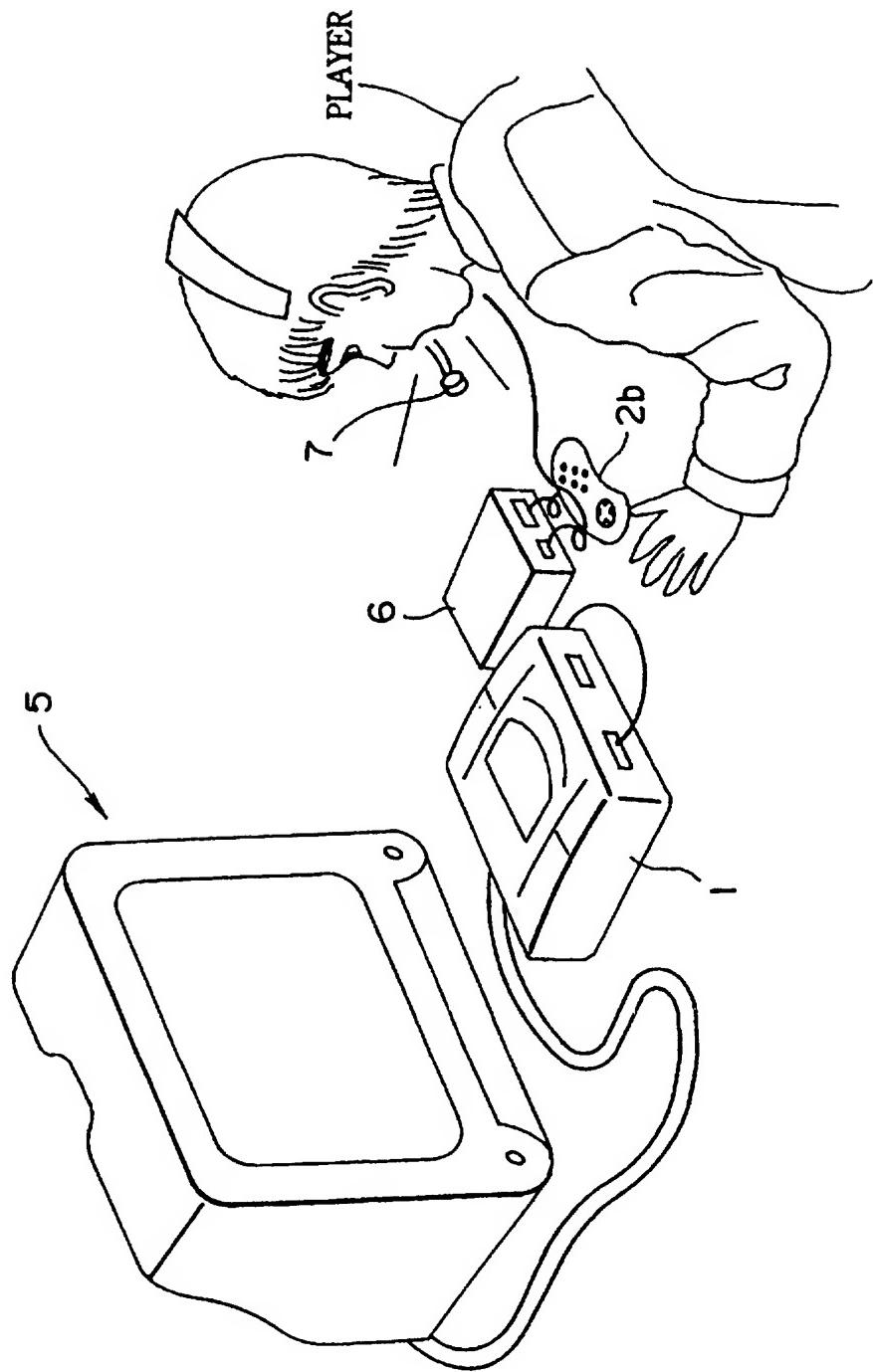


3 / 6

【図 3】

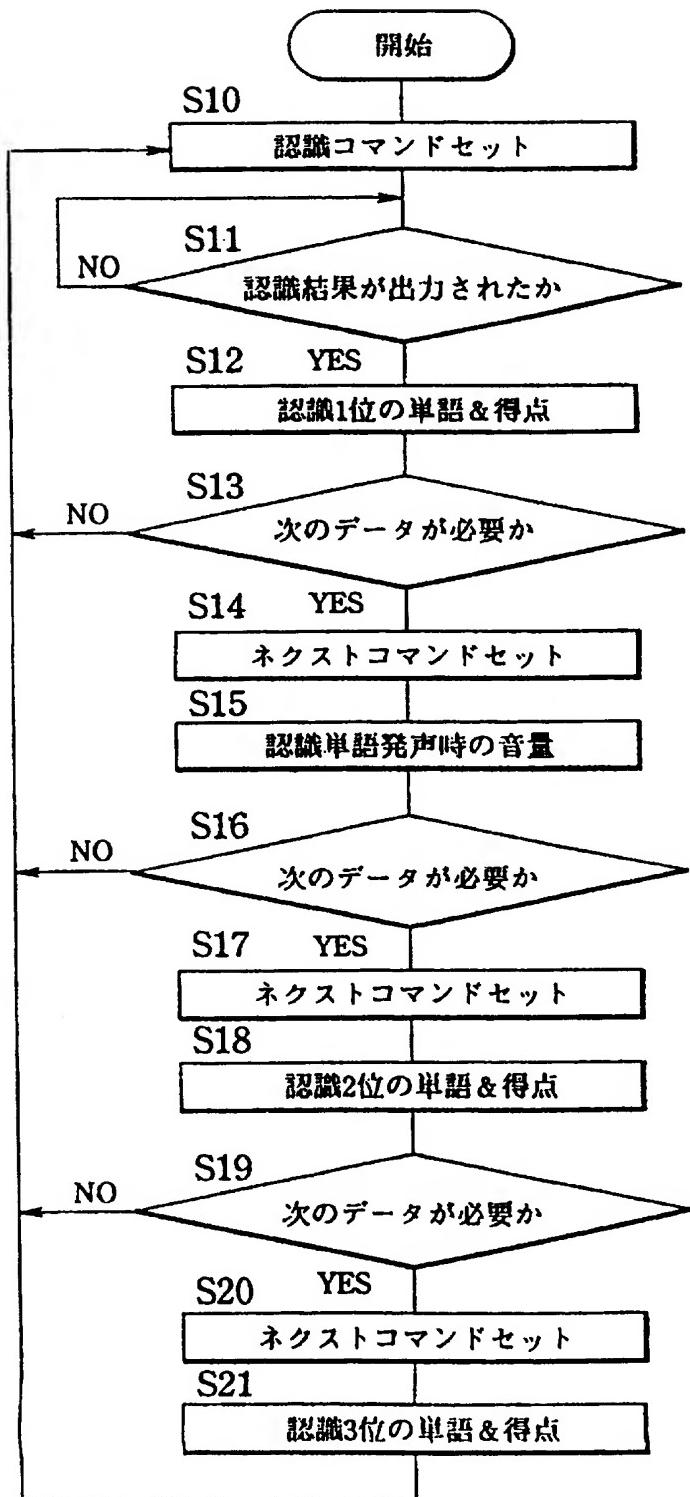


【図 4】



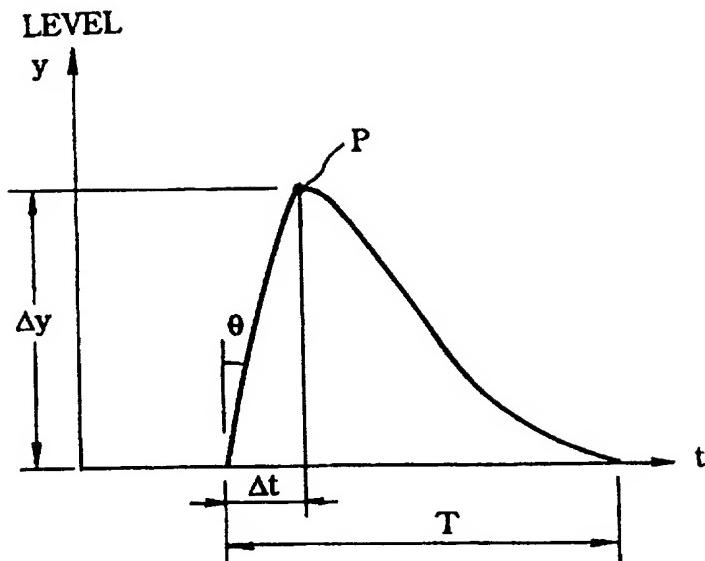
5 / 6

【図 5】

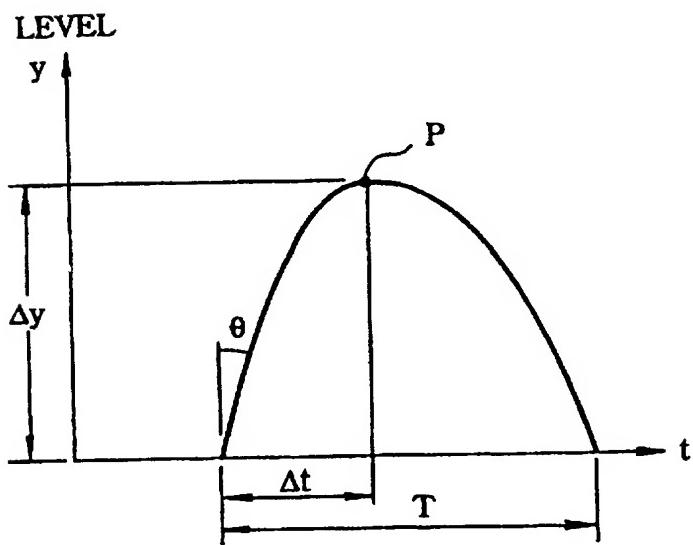


6 / 6

【図 6】



【図 7】



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP97/02421

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl⁶ A63F9/22

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int. Cl⁶ A63F9/22, G10L3/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
 Jitsuyo Shinan Koho 1922 - 1996 Jitsuyo Shinan Toroku
 Kokai Jitsuyo Shinan Koho 1971 - 1997 Koho 1996 - 1997
 Toroku Jitsuyo Shinan Koho 1994 - 1997

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP, 6-315572, A (Ricoh Co., Ltd.), November 15, 1994 (15. 11. 94) (Family: none) Full descriptions; Figs. 1 to 3	1, 4-5, 15-18 2 - 3
Y	Full descriptions; Figs. 1 to 3	6 - 14
A	Full descriptions; Figs. 1 to 3	
A	JP, 64-23299, A (Ricoh Co., Ltd.), January 25, 1989 (25. 01. 89) (Family: none) Full descriptions; Figs. 1 to 4	1 - 18
A	JP, 4-276280, A (Sony Corp.), October 1, 1992 (01. 10. 92) (Family: none) Full descriptions; Figs. 1 to 3	1 - 18
A	Microfilm of the specification and drawings annexed to the written application of Japanese Utility Model Application No. 91363/1988 (Laid-open No. 13587/1990) (Takehiko Kikuchi), January 29, 1990 (29. 01. 90) (Family: none) Full descriptions; Figs. 1, 2	12 - 13



Further documents are listed in the continuation of Box C.



See patent family annex.

• Special categories of cited documents:	
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier document but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
October 3, 1997 (03. 10. 97)Date of mailing of the international search report
October 14, 1997 (14. 10. 97)Name and mailing address of the ISA/
Japanese Patent Office
Facsimile No.Authorized officer
Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP97/02421

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP, 8-123448, A (Sega Enterprises, Ltd.), May 17, 1996 (17. 05. 96) (Family: none) Full descriptions; Figs. 1 to 9	1 - 18
A	JP, 7-222868, A (Sega Enterprises, Ltd.), August 22, 1995 (22. 08. 95) (Family: none) Full descriptions; Figs. 1 to 7	1 - 18
A	JP, 7-175567, A (Nippon Steel Corp.), July 14, 1995 (14. 07. 95) (Family: none) Full descriptions; Figs. 1 to 4	1 - 18

A. 発明の属する分野の分類(国際特許分類(IPC))

Int. C1° A63F9/22

B. 調査を行った分野

調査を行った最小限資料(国際特許分類(IPC))

Int. C1° A63F9/22, G10L3/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-1997年
日本国登録実用新案公報	1994-1997年
日本国実用新案登録公報	1996-1997年

国際調査で使用した電子データベース(データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X	J P, 6-315572, A (株式会社リコー) 15. 11月. 1994 (15. 11. 94) (ファミリーなし) 全文, 第1-3図	1, 4-5, 15-18
Y	全文, 第1-3図	2-3
A	全文, 第1-3図	6-14
A	J P, 64-23299, A (株式会社リコー) 25. 1月. 1989 (25. 01. 89) (ファミリーなし) 全文, 第1-4図	1-18
A	J P, 4-276280, A (ソニー株式会社) 1. 10月. 1992 (01. 10. 92) (ファミリーなし) 全文, 第1-3図	1-18

 C欄の続きにも文献が列挙されている。 パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

- 「A」特に関連のある文献ではなく、一般的技術水準を示すもの
- 「E」先行文献ではあるが、国際出願日以後に公表されたもの
- 「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献(理由を付す)
- 「O」口頭による開示、使用、展示等に言及する文献
- 「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

03. 10. 97

国際調査報告の発送日

14.10.97

国際調査機関の名称及びあて先

日本国特許庁 (ISA/JP)

郵便番号100

東京都千代田区霞が関三丁目4番3号

特許庁審査官(権限のある職員)

2B 9209

植野 孝郎

電話番号 03-3581-1101 内線 3238

国際調査報告

国際出願番号 PCT/JP97/02421

C(続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	日本国実用新案登録出願 63-91363号 (日本国実用新案登録出願公開 2-13587号) の願書に添付した明細書及び図面の内容を撮影したマイクロフィルム (菊池 武彦) 29. 1月. 1990 (29. 01. 90) (ファミリーなし) 全文, 第1-2図	12-13
A	J P, 8-123448, A (株式会社セガ・エンタープライゼス) 17. 5月. 1996 (17. 05. 96) (ファミリーなし) 全文, 第1-9図	1-18
A	J P, 7-222868, A (株式会社セガ・エンタープライゼス) 22. 8月. 1995 (22. 08. 95) (ファミリーなし) 全文, 第1-7図	1-18
A	J P, 7-175567, A (新日本製鐵株式会社) 14. 7月. 1995 (14. 07. 95) (ファミリーなし) 全文, 第1-4図	1-18